Estimating the Parameters of a Continuous-Time Markov Chain from Discrete-Time Data with ctmcd

Marius Pfeuffer
Department of Statistics and Econometrics, University of Erlangen-Nuremberg
July 6, 2017
The Embedding Problem for Markov Chains

Anatomy of Missing Data Situation

- Continuous-time Markov chain (CTMC) model
- Observations only available at discrete-time level (at times 0 and \( T \))

\[
\text{State}(\text{Time}) \quad s(0) \quad s(\tau_1) \quad s(\tau_2) \quad s(\tau_{K-1}) \quad s(\tau_K) \quad s(T)
\]

Parameters of a CTMC

- \( Q \): Generator matrix, intensity matrix, (transition) rate matrix, ...
- Properties: \( q_{ij} \in [0, \infty) \), \( j \neq i \) and \( q_{ii} = -\sum_{j \neq i} q_{ij} \)
- Matrix exponential relationship between conditional discrete-time transition matrices \( P_{\tau_0+\Delta \tau|\tau_0} \) and generator matrix \( Q \)

\[
P_{\tau_0+\Delta \tau|\tau_0} = \exp(Q\Delta \tau)
\]
Standard and Poor’s Corporate Credit Rating Transitions, 2000

**Transition Data for Specific Discrete Time Horizon**

<table>
<thead>
<tr>
<th>From</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>208</td>
<td>22</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AA</td>
<td>5</td>
<td>777</td>
<td>67</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>55</td>
<td>1428</td>
<td>135</td>
<td>6</td>
<td>1</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>BBB</td>
<td>1</td>
<td>6</td>
<td>65</td>
<td>1514</td>
<td>66</td>
<td>9</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>40</td>
<td>886</td>
<td>75</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>5</td>
<td>3</td>
<td>6</td>
<td>48</td>
<td>793</td>
<td>47</td>
<td>53</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>13</td>
<td>77</td>
<td>19</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>To</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0.897</td>
<td>0.095</td>
<td>0.009</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AA</td>
<td>0.006</td>
<td>0.911</td>
<td>0.079</td>
<td>0.005</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>0.034</td>
<td>0.873</td>
<td>0.083</td>
<td>0.004</td>
<td>0.001</td>
<td>0.004</td>
<td>0.002</td>
</tr>
<tr>
<td>BBB</td>
<td>0.011</td>
<td>0.004</td>
<td>0.039</td>
<td>0.907</td>
<td>0.04</td>
<td>0.005</td>
<td>0.002</td>
<td>0.004</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0.004</td>
<td>0.001</td>
<td>0.039</td>
<td>0.87</td>
<td>0.074</td>
<td>0.009</td>
<td>0.003</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0.005</td>
<td>0.003</td>
<td>0.006</td>
<td>0.05</td>
<td>0.83</td>
<td>0.049</td>
<td>0.055</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.009</td>
<td>0.118</td>
<td>0.7</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

**Generator Matrix Estimate**

Transition Predictions for Arbitrary Discrete Time Horizons
Matrix Logarithm Adjustment Approaches

Candidate generator matrix

\[ \tilde{Q} = \frac{\log(P_T|0)}{T} \]

→ Negative off-diagonal elements can occur

Diagonal adjustment (Israel et al., 2001)

\[ \hat{q}_{ij} = \begin{cases} 0 & \text{if } i \neq j \land \tilde{q}_{ij} < 0 \\ \tilde{q}_{ij} & \text{else} \end{cases} \quad \text{and} \quad \hat{q}_{ii} = -\sum_{j \neq i} \hat{q}_{ij} \]

Weighted adjustment (Israel et al., 2001)

\[ \hat{q}_{ij} = \begin{cases} 0 & \text{if } i \neq j \land \tilde{q}_{ij} < 0 \\ \tilde{q}_{ij} + \frac{|\tilde{q}_{ij}|}{\sum_{j \neq i} |\tilde{q}_{ij}|} \sum_{j \neq i, \tilde{q}_{ij} < 0} \tilde{q}_{ij} & \text{else} \end{cases} \]

Quasi-optimization (Kreinin and Sidelnikova, 2001)

\[ \tilde{Q} = \arg\min_{Q} \|Q - \tilde{Q}\|^2 \]
Matrix Logarithm Adjustment Approaches

Access methods in ctmcd:

gmda <- gm(tm=tm_rel, te=1, method="DA")
gmwa <- gm(tm=tm_rel, te=1, method="WA")
gmqo <- gm(tm=tm_rel, te=1, method="QO")

Plot results:

plot(gmqo)
Maximum Likelihood Estimation

Expectation-maximization algorithm

- Iterative imputation of missing continuous time data
  - $R_i(T)$: Cumulative continuous-time sojourn times
  - $N_{ij}(T)$: Cumulative continuous-time state changes
- Monotone optimization of likelihood function

EM-Algorithm (Bladt and Sørensen, 2005)

- initialization
  set $\tilde{Q}^{(0)}$
  for $n = 1$ to $n_{\text{iter}}$ do
    ① E-step
    derive $\tilde{R}_i(T) = \mathbb{E}(R_i(T)|\tilde{Q}^{(n-1)}, s(0) = s_0, s(T) = s_T)$
    derive $\tilde{N}_{ij}(T) = \mathbb{E}(N_{ij}(T)|\tilde{Q}^{(n-1)}, s(0) = s_0, s(T) = s_T)$
    ② M-step
    derive $\hat{q}_{ij}^{(n)} = \frac{\tilde{N}_{ij}(T)}{\tilde{R}_i(T)}$
  end for
E-Step

Conditional Expectations

\[ E(R_i(T)|\tilde{Q}, s(0) = s_0, s(T) = s_T) = \frac{u_{s_0}^T \left( \int_0^T \exp(\tilde{Q}(s))u_i u_i^T \exp(\tilde{Q}(T - s))ds \right) u_{s_T}}{u_{s_0}^T \exp(\tilde{Q}T)u_{s_T}} \]

\[ E(N_{ij}(T)|\tilde{Q}, s(0) = s_0, s(T) = s_T) = \frac{u_{s_0}^T \tilde{q}_{ij} \left( \int_0^T \exp(\tilde{Q}(s))u_i u_j^T \exp(\tilde{Q}(T - s))ds \right) u_{s_T}}{u_{s_0}^T \exp(\tilde{Q}T)u_{s_T}} \]

Implementation

- Time consuming evaluation of matrix exponential function → classical numerical integration techniques fail
- Approach of van Loan, 1978 / Inamura, 2006
Expectation-Maximization Algorithm

Initialization:

\[
\begin{align*}
gm0 & \leftarrow \text{matrix}(1, 8, 8) \\
\text{diag}(gm0) & \leftarrow 0 \\
\text{diag}(gm0) & \leftarrow -\text{rowSums}(gm0) \\
gm0[8,] & \leftarrow 0
\end{align*}
\]

Access method in ctmcd:

\[
\begin{align*}
gmem & \leftarrow \text{gm(tm=tm\_abs,te=1,method="EM", gmguess=gm0)} \\
\text{plot(gmem)}
\end{align*}
\]
Confidence Interval

Method of Oakes, 1999/Bladt and Sørensen, 2009
- Asymptotic normality of maximum likelihood estimate
- Computation only possible for intensity estimates of certain minimum size for numerical reasons

Access method in ctmcd:
```r
ciem <- gmci(gmem,alpha=0.05)
plot(ciem)
```

95% Wald Confidence Interval (Oakes Standard Error)
Gibbs Sampler

Posterior Distribution

- CTMC likelihood function: \( L(Q|N_{ij}(T), R_i(T')) \propto \prod_{i=1}^{I-1} \prod_{j \neq i} q_{ij}^{N_{ij}(T)} \exp(-q_{ij} R_i(T')) \)
- Conjugate gamma prior: \( f(q_{ij}) \propto q_{ij}^{\alpha - 1} \exp(-q_{ij} \beta) \)

Gibbs Sampler (Bladt and SØrensen, 2005)

▷ initialization
set \( \tilde{Q}^{(0)} \)

for \( n = 1 \) to \( n_{\text{burn-in}} + n_{\text{iter}} \) do
  ▷ draw from full conditional distributions
  simulate \( s(t)|\tilde{Q}^{(n-1)}, s(0) = s_0, s(T) = s_T \) and derive \( \tilde{N}_{ij}(T) \) and \( \tilde{R}_i(T) \)
  draw \( \tilde{q}_{ij}^{(n)} \) from \( \Gamma(N_{ij}(T) + \alpha_{ij}, R_i(T) + \beta_i) \)
end for

derive \( \hat{q}_{ij} = \sum_{n=n_{\text{burn-in}}+1}^{n_{\text{iter}}} \tilde{q}_{ij}^{(n)} \)

Endpoint-Conditioned CTMC Sampling

- Modified rejection sampling scheme (Nielsen, 2002)
- Uniformization sampling approach (Fearnhead and Sherlock, 2006)
Gibbs Sampler

Prior parametrization:

```r
pr <- list()
pr[[1]] <- matrix(1,8,8)
pr[[1]][8,] <- 0
pr[[2]] <- c(rep(5,7),Inf)
```

Access method in ctmcd:

```r
gmgs <- gm(tm=tm_abs,te=1,method="GS",
burnin=100,prior=pr)
plot(gmgs)
```
Credibility Interval

Approach of Bladt and Sørensen, 2009
▶ Equal-tailed credibility interval
▶ Empirical quantiles of Gibbs sampler draws

Access method in `ctmcd`:
```r
cigs <- gmci(gmgs, alpha=0.05)
plot(cigs)
```

95% Equal Tailed Credibility Interval

<table>
<thead>
<tr>
<th>From</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>[−0.19; −0.095]</td>
<td>[0.066; 0.16]</td>
<td>[0.00018; 0.027]</td>
<td>[9.9e−05; 0.017]</td>
<td>[9.3e−05; 0.016]</td>
<td>[0.00012; 0.017]</td>
<td>[0.00016; 0.015]</td>
<td>[9.8e−05; 0.014]</td>
</tr>
<tr>
<td>AA</td>
<td>[0.0027; 0.014]</td>
<td>[−0.13; −0.082]</td>
<td>[0.067; 0.11]</td>
<td>[8.3e−05; 0.0095]</td>
<td>[3e−05; 0.0047]</td>
<td>[2.2e−05; 0.0043]</td>
<td>[4.4e−05; 0.0045]</td>
<td>[4.3e−05; 0.0042]</td>
</tr>
<tr>
<td>A</td>
<td>[1.8e−05; 0.0024]</td>
<td>[0.028; 0.048]</td>
<td>[−0.16; −0.12]</td>
<td>[0.078; 0.11]</td>
<td>[0.00027; 0.0067]</td>
<td>[4.4e−05; 0.0035]</td>
<td>[0.0019; 0.0095]</td>
<td>[0.00033; 0.0057]</td>
</tr>
<tr>
<td>BBB</td>
<td>[7e−05; 0.0036]</td>
<td>[0.00082; 0.0072]</td>
<td>[0.034; 0.054]</td>
<td>[−0.12; −0.089]</td>
<td>[0.035; 0.056]</td>
<td>[0.001; 0.0092]</td>
<td>[0.00039; 0.0065]</td>
<td>[0.0011; 0.0075]</td>
</tr>
<tr>
<td>BB</td>
<td>[2.4e−05; 0.0038]</td>
<td>[0.0011; 0.01]</td>
<td>[5.9e−05; 0.0054]</td>
<td>[0.031; 0.059]</td>
<td>[−0.18; −0.12]</td>
<td>[0.068; 0.11]</td>
<td>[0.0026; 0.017]</td>
<td>[8.8e−05; 0.0066]</td>
</tr>
<tr>
<td>B</td>
<td>[3.5e−05; 0.0047]</td>
<td>[0.0019; 0.013]</td>
<td>[0.00076; 0.0093]</td>
<td>[0.0011; 0.015]</td>
<td>[0.044; 0.077]</td>
<td>[−0.23; −0.17]</td>
<td>[0.049; 0.09]</td>
<td>[0.038; 0.071]</td>
</tr>
<tr>
<td>C</td>
<td>[0.00026; 0.027]</td>
<td>[0.00029; 0.038]</td>
<td>[0.00026; 0.037]</td>
<td>[2e−04; 0.04]</td>
<td>[0.001; 0.058]</td>
<td>[0.084; 0.25]</td>
<td>[−0.55; −0.3]</td>
<td>[0.12; 0.31]</td>
</tr>
<tr>
<td>D</td>
<td>[0; 0]</td>
<td>[0; 0]</td>
<td>[0; 0]</td>
<td>[0; 0]</td>
<td>[0; 0]</td>
<td>[0; 0]</td>
<td>[0; 0]</td>
<td>[0; 0]</td>
</tr>
</tbody>
</table>

To
Credit Risk Modeling

- Upcoming accounting regulation: IFRS 9
- Takes effect Jan 1, 2018
- Requires evaluation of credit risk for whole maturity of an institute’s credits
- Banks often only have discrete-time rating transition data available
References

Methods


Packages

References

EM Algorithm/Gibbs Sampler Methodology

Confidence/Credibility Intervals